Deep Learning - Week 5

1. Which of the following is the most appropriate description of the method used in PCA to achieve dimensionality reduction?

   (a) PCA achieves this by discarding a random subset of features in the dataset

   (b) PCA achieves this by selecting those features in the dataset along which the variance of the dataset is maximised

   (c) PCA achieves this by retaining the those features in the dataset along which the variance of the dataset is minimised

   (d) PCA achieves this by looking for those directions in the feature space along which the variance of the dataset is maximised

   Correct Answer: (d)
   Solution: PCA looks for a new set of directions in feature space such that the first few directions capture the maximum variance in the data. It does this by re-orienting the feature axes, which can be thought of as rotating the axes in the feature space.

2. What is/are the limitations of PCA?

   (a) It can only identify linear relationships in the data.

   (b) It can be sensitive to outliers in the data.

   (c) It is computationally less efficient than autoencoders

   (d) It can only reduce the dimensionality of a dataset by a fixed amount.

   Correct Answer: (a),(b)
   Solution: PCA can be sensitive to outliers in the data, since the principal components are calculated based on the covariance matrix of the data. Outliers can have a large impact on the covariance matrix and can skew the results of the PCA. Also, it can only capture linear relationships in the data.

3. The following are possible numbers of linearly independent eigenvectors for a 7×7 matrix. Choose the incorrect option.

   (a) 1

   (b) 3

   (c) 9

   (d) 5

   (e) 8

   Correct Answer: (c),(e)
   Solution: A $n \times n$ matrix can have between 1 and n linearly independent eigenvectors.

4. Find the singular values of the following matrix: $\begin{bmatrix} -4 & -6 \\ 3 & -8 \end{bmatrix}$

(a) $\sigma_1 = 10, \sigma_2 = 5$

(b) $\sigma_1 = 1, \sigma_2 = 0$

(c) $\sigma_1 = 100, \sigma_2 = 25$

(d) $\sigma_1 = \sigma_2 = 0$

Correct Answer: (a)

Solution: Let $A = \begin{bmatrix} -4 & -6 \\ 3 & -8 \end{bmatrix}$. Then,

$$A^T A = \begin{bmatrix} 25 & 0 \\ 0 & 100 \end{bmatrix}.$$

The singular values of A are the positive square roots of the eigenvalues of $A^T A$. Therefore,

$$\sigma_1 = 10 \quad \text{and} \quad \sigma_2 = 5.$$

5. PCA is performed on a mean-centred dataset in $\mathbb{R}^3$ If the first principal component is $\frac{1}{\sqrt{6}}(1, -1, 2)$, which of the following could be the second principal component?

(a) $(1, -1, 2)$

(b) $(0, 0, 0)$

(c) $\frac{1}{\sqrt{5}}(0, 1, 2)$

(d) $\frac{1}{\sqrt{2}}(-1, -1, 0)$

Correct Answer: (d)

Solution: The principal components are orthogonal eigenvectors of the covariance matrix. Since they are eigenvectors, $(0, 0, 0)$ is ruled out, A zero vector cannot be a principal component because it has no direction. Since they have to be orthogonal, only the option d is correct.

Questions 6-9 are based on common data.

Consider the following data points x_1, x_2, x_3 to answer following questions: $x_1 = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$, $x_2 = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$, $x_3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

6. What is the mean of the given data points x_1, x_2, x_3?

(a) $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

(b) $\begin{bmatrix} 1.67 \\ 1.67 \end{bmatrix}$

(c) $\begin{bmatrix} 2 \\ 2 \end{bmatrix}$

(d) $\begin{bmatrix} 0.33 \\ 0.33 \end{bmatrix}$

Correct Answer: d)

Solution: Mean of $x_1, x_2, x_3 = \frac{x_1+x_2+x_3}{3} = \frac{1}{3}\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}0.33\\0.33\end{bmatrix}$

7. The covariance matrix $C = \frac{1}{n}\sum_{i=1}^{n}(x-\bar{x})(x-\bar{x})^T$ is given by: ($\bar{x}$ is mean of the data points)

   (a) $\begin{bmatrix}8.66 & -7.33\\-7.33 & 8.66\end{bmatrix}$

   (b) $\begin{bmatrix}2.88 & -2.44\\-2.44 & 2.88\end{bmatrix}$

   (c) $\begin{bmatrix}0.22 & -0.22\\-0.22 & 0.22\end{bmatrix}$

   (d) $\begin{bmatrix}5.33 & -0.33\\-5.33 & 0.33\end{bmatrix}$

   Correct Answer: b)

   Solution: $\bar{x} = \begin{bmatrix}0.3333\\0.3333\end{bmatrix}$ $x_1 - \bar{x} = \begin{bmatrix}-2.33333\\1.66667\end{bmatrix}$ $x_2 - \bar{x} = \begin{bmatrix}1.66667\\-2.33333\end{bmatrix}$ $x_3 - \bar{x} = \begin{bmatrix}0.666667\\0.666667\end{bmatrix}$

   Now, let's calculate $(x-\bar{x})(x-\bar{x})^T$ for each point:

   For x_1: $\begin{bmatrix}-2.33333\\1.66667\end{bmatrix}\begin{bmatrix}-2.33333 & 1.66667\end{bmatrix} = \begin{bmatrix}5.44444 & -3.88889\\-3.88889 & 2.77778\end{bmatrix}$

   For x_2: $\begin{bmatrix}1.66667\\-2.33333\end{bmatrix}\begin{bmatrix}1.66667 & -2.33333\end{bmatrix} = \begin{bmatrix}2.77778 & -3.88889\\-3.88889 & 5.44444\end{bmatrix}$

   For x_3: $\begin{bmatrix}0.666667\\0.666667\end{bmatrix}\begin{bmatrix}0.666667 & 0.666667\end{bmatrix} = \begin{bmatrix}0.444444 & 0.444444\\0.444444 & 0.444444\end{bmatrix}$

   Sum these matrices: $\sum_{i=1}^{n}(x-\bar{x})(x-\bar{x})^T$

Now, multiply by $\frac{1}{n} = \frac{1}{3}$:

$C = \begin{bmatrix}2.88889 & -2.44444\\-2.44444 & 2.88889\end{bmatrix}$

Therefore, the correct covariance matrix is $\begin{bmatrix}2.88889 & -2.44444\\-2.44444 & 2.88889\end{bmatrix}$

8. The maximum eigenvalue of the covariance matrix C is:

   (a) 1

   (b) 5.33

   (c) 0.44

   (d) 0.5

Correct Answer: b)

Solution: $C = \begin{bmatrix} 2.88889 & -2.44444 \\ -2.44444 & 2.88889 \end{bmatrix}$

This gives us two eigenvalues: $\lambda_1 = 0.44445$ $\lambda_2 = 5.33333$

The maximum eigenvalue is $\lambda_2 = 5.33$.

9. The eigenvector corresponding to the maximum eigenvalue of the given matrix C is:

   (a) $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

   (b) $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$

   (c) $\begin{bmatrix} 0.67 \\ 0 \end{bmatrix}$

   (d) $\begin{bmatrix} -1.48 \\ 1 \end{bmatrix}$

   Correct Answer: b)

   Solution: Using the maximum eigenvalue found earlier, we solve the equation $(C - \lambda I)v = 0$ to find the eigenvector v. The eigenvector corresponding to the maximum eigenvalue is $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

10. Given that A is a 2×2 matrix, what is the determinant of A, if its eigenvalues are 6 and 7?

    Correct Answer: 42

    Solution: The determinant of a matrix is defined as the product of its eigenvalues. Therefore, if a matrix has eigenvalues $\lambda 1$ and $\lambda 2$, its determinant is given by $det(A) = \lambda 1 * \lambda 2$.